Student book answers

Chapter 8 Data management

Test your knowledge

1 The following answers given are example answers only. Other answers may also be acceptable.

a Financial and contact details for the customers, recording financial transactions
b Keeping track of current financial (paid-up) members
c Storing information on customer addresses, credit card details, or purchasing history.
d Storing and tallying votes

2 Metadata is information about the data within the database. Metadata allows databases to locate files more easily if they match certain search criteria.

3 Any three of the following are acceptable.

- Reduced data redundancy through the centralisation of different data sources into the one database file.
- Data consistency as a flow-on effect of reduced redundancy
- Sharing of data allows multiple users to access the same data simultaneously. For example, an accounts department might want to send an invoice to a customer, while the marketing department wants to send an invitation to a special product launch. Both departments will send their emails addressed the same way and including the same personal details because they are both taking the data from the one database in their company.
- Maintaining data integrity via the use of validation
- Data security measures mean only authorised users can access certain data. Data is often 'read only' to most users to restrict who can edit it.
- Multiple data elements can be updated in different tables when a transaction takes place.
- A common structured query language or SQL is used in all packages to select data.
- Input forms can be created to make data entry easier for users.
- Reports can be created to display information extracted from database files. Reports can be saved and modified for later reuse. Reports can include graphic representations like charts.
- Concurrency of data is maintained when two users try to edit the same record simultaneously.
- Data is independent of the database application so properties of fields can be modified without altering the application or the database file itself.

4 List three limitations of a DBMS.

Answer: Any three of the following.

- Equipment (software and hardware) can be costly, especially if upgrades are needed on a regular basis.
- Conversion from a computer file used for data storage, like a spreadsheet, to a DBMS can be costly and time consuming. Fields and tables must be carefully constructed to ensure that no data is lost.
- Staff need training to use the new system. This can be costly.
- Specialised database administrators need to be employed.

- If a database is corrupted, say due to a technical failure, it has a huge impact across the entire organisation.

5 What are three benefits of using a DBMS?

Answer: Any three of the following.

- It allows all of your relevant data to be stored in one place. This can reduce the amount of time and effort that would be involved in manually searching for information.
- An increase in speed in finding information could also save you money.
- When you need to ask a particular question of your data, it can be answered in a timely manner with both relevance and completeness.
- Validation and calculation formulas can be used to guarantee accuracy of data and information respectively.
- Easy to use interface and selection tools can make your data more accessible and usable.
- Selected data can be formatted into well-presented and attractive reports that make the information clearer and more readable to a wider audience on a variety of different platforms. These all contribute getting across the message that the information needs to convey to the end user.

6 What are three risks associated with using a DBMS?

Answer: Any three of the following.

- Garbage in equals garbage out (GIGO). Input must be validated to avoid inaccurate data being entered.
- Privacy must be protected.
- Security must be protected.
- Database systems can be complex, difficult, and time-consuming to design.
- Training is required for all programmers and users or data can be damaged.
- Costs of purchasing suitable hardware and software can be high.

7 Define each of the following DBMS terms.

a Field

Answer: Individual data elements.

b Table

Answer: A table holds a set of records. Each column represents a field and each row represents a record.

c Primary key

Answer: A field containing a unique value for a record which enables it to be distinguished from other records.

d Input form

Answer: A form that allows user to enter data into a database via a more user friendly interface than a table.

e Query

Answer: The database object which defines the selection of specific sets of data.

f SQL

Answer: Structured query language: A language used by many database packages to provide a consistent way of selecting required data from a table.

g Report

Answer: A report formats and presents the data selected by a query as usable information so that the reader can more clearly understand it.

h Macro

Answer: A series of steps that have been automated so that a single button or key combination can be used to complete the whole process.

i File

Answer: A collection of records.

8 There are different data types that can be used in a DBMS. In the table below, write the most appropriate data type and an example of data beside each field.

Answer:

Field	Data type	Example
Company name	Text	Google
Australian postcode	Text	3000
Mobile phone number	Text	0419123456
Height in cm	Numeric	156
Preferred pizza size: S, M or L	Character	M
Would you like to be sent our newsletter?	Boolean	Yes
Date of dispatch	Date/Time	01/01/2016

9 Explain how the three validation checks could be used on a field that records people's weight in kilograms.

Answer: Existence check – is there a value in the field?
Data type check – is the value in the field numeric?
Range check – does the value fit within the acceptable range (e.g. between 30 and 200&&KG).

10 What are two characteristics of an effective input form?

Answer: Students' answers will vary but may include:

- logical arrangement
- mandatory entry fields
- labels showing examples of expected values
- drop down lists with prefilled options
- suggested spelling

- good design, e.g. adapting the fields based on previous choices.

11 What does the term 'user experience' mean?

Answer: User experience refers to how a user interacts with a system, especially in regards to how easy and pleasant it is for the user.

12 Explain how the file naming convention for database objects works.

Answer: Every object should have a prefix which identifies their object type. The rest of the name should be something that is easily identified by a developer or a user. Reports and forms should be very descriptive as they are often interacted with. Fields should have a prefix which identifies the object they are part of.

13 How does a data dictionary assist in the design of a database?

Answer: A data dictionary is a specific type of metadata that records the details of structural elements in a database file. It includes information about the size and data type of fields as well as validation rules.

14 How does the purpose of a layout diagram differ from a data dictionary?

Answer: A layout diagram is a design tool that helps plan the appearance of input forms and reports and shows the interface developer how to build the product. For example, it details the location and format of headings, labels and fields. A data dictionary, on the other hand, is metadata containing these structural elements that sit within the database.

15 Recall the purpose of query criteria.

Answer: It enables selection of a specific set of data.

16 Provide reasons why is it important to sort information in a report.

Answer: It increases effectiveness because it is easier to find relevant information.

17 How do summary totals in reports assist with effectiveness of a solution?

Answer: They increase readability and communication of the message.

18 Recall the design tools that are commonly used to plan a macro.

Answer: Pseudocode (structured English) or a flowchart.

19 State the role of input devices.

Answer: Input devices speed up data entry.

20 How does data input by a keyboard differ to that of a bar code reader?

Answer: A user has to manually type in data with a keyboard and speed and accuracy depends on the user. A bar code reader on the other hand inputs data quickly and accurately via an optical reader.

21 What is an advantage of the use of a wireless data collection device in a supermarket?

Answer: You can take it around the store to the objects for sale.

22 List three differences between internal solid state drives and removable flash drives.

Answer: The storage chips on solid state drives are usually faster and more expensive. Removable drives are limited by the speed of the connection interface they use (for example, USB). Removable drives are also more prone to damage or loss.

23 How can RAID help to protect stored data?

Answer: RAID (redundant array of inexpensive disks) are banks of hard drives that are configured in such a way that if one or two fail the data they contain isn't lost. Instead the remaining disks can rebuild the missing data in their own storage areas.

24 Why is communication of data critical for many databases?

Answer: They are often networked via internal networks or over the internet in order to share their data with all users who may be in different locations. The data must be safeguarded with encryption during communication.

25 What is the difference between a softcopy and a hardcopy? Provide an example of where each may be more appropriate.

Answer: Softcopy is information displayed on an electronic device, like a screen, and is temporary. A hardcopy is a physical object that has been printed out and is more permanent. You would normally only read your emails onscreen but you might print it out if it contained a concert ticket, for example.

26 How does a deliberate threat to data differ from an accidental threat?

Answer: A deliberate threat by a hacker or malware specifically intends to alter, delete or steal data. Accidental threats, such as data being lost or altered by mistake, are not done on purpose and are not malicious.

27 Describe what is understood as an event based threat to data.

Answer: Data loss or corruption occurring from power failure or a crash of a storage device. Another example is a remote connection failure which might result in an incomplete record being saved.

28 How can backing up help to protect the security of data?

Answer: You have another copy of the data in case the original gets lost or corrupted. You might also need to go back to a previous version if changes have been made to the 'live' copy that you don't want to keep.

29 What types of electrical protection can be provided by a UPS?

Answer: UPS or Uninterruptable Power Supply regulates electricity coming into a computer and supplies backup electricity for a short duration during a blackout.

30 Why is a combination of username and password used to protect a database, rather than an individual username by itself?

Answer: The password is private where as a username might be more likely to be known by more than one person. Even if an unauthorised person obtains a username they will still need to get the password right to gain access.

31 What is a definition of malware?

Answer: Malicious software that deliberately attacks your computer system.

32 How does encryption secure a data connection between a user inputting data and the database itself?

Answer: Encryption involves coding data into a form that only authorised users can read and decrypt. The encryption and decryption process both require a 'key'. For example, internet transactions will often use SSL or TLS protocols over an HTTPS connection which creates a secure 'pipeline' between the user and the database so that the data transmitted cannot be read without the key.

Apply your knowledge

DATABASE ACTIVITY

The two reports Catherine would like are as follows.

a A report that shows all the debits from her cheque account for the month of November 2016. It should be grouped by transaction category and sorted by transaction date within each category. Be sure to include summary statistics that show the percentage and sub totals for each group as well as an overall total so Catherine can see what money remained at the end of the month.

b A report that lists all expenses from all of Catherine's accounts for the 'Car/Travel' category for the month of November 2016. The report should be sorted by transaction date (from earliest to most recent) and then by amount (smallest to largest amount).

1 Table 8.9 shows the set of data that Catherine has provided you with.

2 Looking carefully at the data, create a data dictionary to plan the database table. You may rearrange the data into different fields as you see fit. Also remember to include validation and to choose appropriate data types and formats.

3 Design a suitable data input form that will enable Catherine to add, review and edit her bank records. Be sure to make use of labels and drop down lists to assist in more accurate data entry.

4 Create designs for the two queries that will be used to create the reports outlined above.

5 Design the two reports outlined above.

6 Design macros that will open, print and close each of the reports when they are run.

7 Using a DBMS package, create the Catherine's bank transactions database. This will involve constructing the table, the input form, the two queries, two reports and the two macros you designed above. Remember to include the data validation and appropriate formulas and functions.

8 Create annotated screenshots of the various database objects you have created to show their key features and how you have met Catherine's needs.

9 In a paragraph, list the roles, functions and characteristics of suitable hardware components that Catherine could use to input, store, communicate and output the data and information in her database.

Answer: Students' answers will vary. The following table has been provided as a guide.

Hardware	Role	Function	Characteristics
Input Keyboard Mouse Touch screen	Manual input of swimmers' data	Users press keys to input data or press buttons to select data to input from premade lists	From 101–105 keys Assignable special function keys Can handle rapid input of typed data 2–3 buttons. Slow data speed
Storage Internal HDD/SSD Cloud	Permanent storage of swimmers' data	Holds data files	Capacity measured in Gigabytes Can hold any type of data files Permanent

Communication Wi-fi	Wireless transmission of data from tablets to club server and cloud	Enables any wi-fi equipped device to share data	Fast transfer speeds Wireless Can be encrypted
Output Screen (softcopy) Print (hardcopy)	Displays information to Felix and treasurer so it can be read and interpreted	Information is displayed on an electronic device Information is printed onto physical media like paper.	More permanent Exists physically Can use a range of media

10 Identify one each of an accidental, deliberate and an event based security threat to the data and information stored in Catherine's database.

Answer: Accidental: Catherine types in the wrong figure.
Deliberate: the computer is hacked.
Event based: the computer crashes and the files are lost.

11 Identify and explain suitable physical and software controls that will protect the data stored in Catherine's database or when it is transmitted (both physically and electronically).

Answer: Controls include backing up the database, providing electrical protection, using usernames and passwords, installing system security software and using encryption.

12 Evaluate the potential success of the DBMS you have created for Catherine by comparing relevant benefits and risks from its use.

Answer: Benefits – all the swimmers' data is stored in the one central location, which saves time and effort. The queries allow timely searching and selecting of relevant data. Swimmers' details are validated when input, which increases accuracy. Reports are clearly laid out which makes the information more readable.
Risks – it is possible that unreasonable data can be entered. There is the possibility of privacy breaches if unauthorised users access the database. Coaches and club committee members will need training.

AUTOMATIC NUMBERPLATE RECOGNITION

1 Identify one each of an Input, Storage, Communication and Output device that might be used in the ANPR system.

Answer: Input: optical scanner
Storage: hard drive on the server
Communication: wireless
Output: computer screen

2 Can you suggest two reasons as to why privacy might be a concern with this system?

Answer: It creates a record for police to monitor people's day-to-day movements and then it stores this information for a long period of time.

3 What might be some other ways the data could be misused?

Answer: It might be used for purposes other than solving crimes, it might be used by agencies other than the police, it might be hacked and the information used for malicious purposes by an unknown party.

4 Suggest two threats that the ANPR database might face.

Answer: A deliberate threat from a hacker deleting, changing or stealing the information or an event based threat occurring when transmission of data is interrupted or the server goes down.

5 Explain a suitable control measure to protect the ANPR data once it is stored in the database and while it is transmitted to and from a police car.

Answer: The plain text could be converted to ciphertext with a suitable encryption method. When transmitting, SSL or TLS over a HTTPS connection can be used.